

## Visual display of objects in field of view for man-machine communication by acquiring information input by user using signal or pattern recognition

**Patent number:** DE10056291

**Publication date:** 2002-05-23

**Inventor:** FEIL HENRY (DE)

**Applicant:** SIEMENS AG (DE)

**Classification:**

- international: G06F3/00

- european: G06F3/00B8, G06F3/00B8D, G06F3/00B8G,  
G06F3/033A1E

**Application number:** DE20001056291 20001114

**Priority number(s):** DE20001056291 20001114

**Also published as:**



WO0241069 (A1)

### Abstract of DE10056291

The method involves manipulating the characteristics and/or actions of at least one displayed object using interactive control commands from a user, detected using sensors etc. The user may be able to navigate in a displayed screen. The information input by the user is acquired and processed using signal- or pattern recognition. No additional manually operable mechanical or touch-sensitive hardware devices are required for input of control commands. A virtual retinal display or head-mounted display may be used to display a virtual environment.

---

Data supplied from the **esp@cenet** database - Worldwide

**THIS PAGE BLANK (USPTO)**

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
23. Mai 2002 (23.05.2002)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 02/41069 A1

(51) Internationale Patentklassifikation<sup>7</sup>: G02B 27/01,  
G06F 3/00

(21) Internationales Aktenzeichen: PCT/DE01/04267

(22) Internationales Anmeldedatum:  
14. November 2001 (14.11.2001)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:  
100 56 291.4 14. November 2000 (14.11.2000) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von  
US): SIEMENS AKTIENGESELLSCHAFT [DE/DE];  
Wittelsbacherplatz 2, 80333 München (DE).

(72) Erfinder; und  
(75) Erfinder/Anmelder (nur für US): FEIL, Henry [DE/DE];  
Grünwalderstrasse 224, 81545 München (DE).

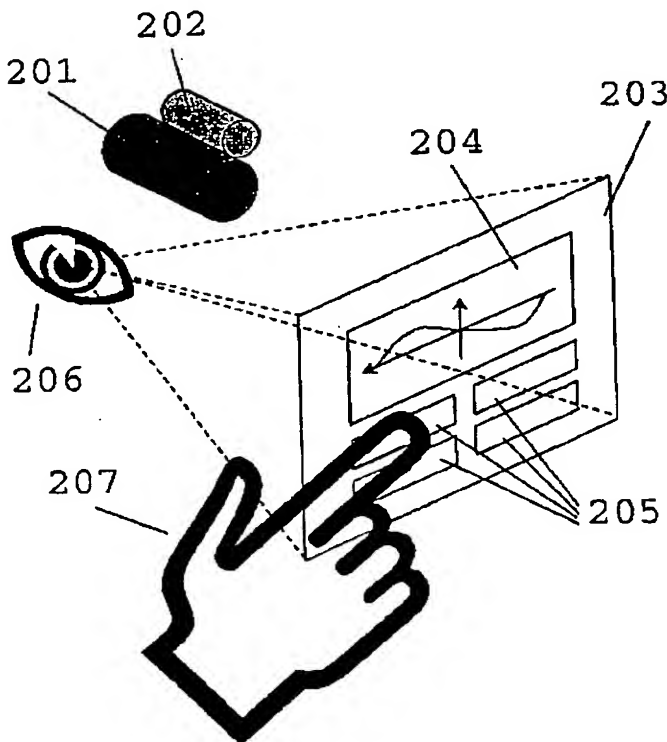
(74) Gemeinsamer Vertreter: SIEMENS AKTIENGE-  
SELLSCHAFT; Postfach 22 16 34, 80506 München  
(DE).

(81) Bestimmungsstaaten (national): CN, US.

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR VISUALLY REPRESENTING AND INTERACTIVELY CONTROLLING VIRTUAL OBJECTS ON  
AN OUTPUT VISUAL FIELD

(54) Bezeichnung: VERFAHREN ZUR VISUELLEN DARSTELLUNG UND INTERAKTIVEN STEUERUNG VON VIRTUEL-  
LEN OBJEKTEN AUF EINEM AUSGABE-SICHTFELD



(57) Abstract: The invention relates to a method for graphically visualizing objects by means of which the objects represented on at least one visual field (203) are comfortably and reliably manipulated, controlled or influenced, with regard to their features and/or actions, by interactive control commands of a user. Technologies used for inputting the control commands issued by the user can, in addition to standard manually operable mechanical or touch-sensitive input mechanisms such as a keyboard, a mouse, a trackball, a joystick, a graphics tablet and stylus, tactile displays, etc., involve the use of devices (501, 502 and 503) for recording, recognizing, interpreting and processing acoustic and/or optical signals of a user. The user is thus no longer dependent upon the presence of additional hardware devices for manually inputting control commands. The evaluation of the input information can instead or additionally ensue by using methods involving signal or pattern recognition. This inventive step enables a conversion of a simple output visual field, with which the user is not provided with any ability to control, into an interactively operable input and output visual field (203). The type of input method can be individually adapted to the abilities of the user.

[Fortsetzung auf der nächsten Seite]

WO 02/41069 A1

**THIS PAGE BLANK (USPTO)**



①9 BUNDESREPUBLIK  
DEUTSCHLAND



DEUTSCHES  
PATENT- UND  
MARKENAMT

⑫ **Offenlegungsschrift**  
⑩ **DE 100 56 291 A 1**

⑤1 Int. Cl.<sup>7</sup>:  
**G 06 F 3/00**

②1 Aktenzeichen: 100 56 291.4  
②2 Anmeldetag: 14. 11. 2000  
④3 Offenlegungstag: 23. 5. 2002

DE 100 56 291 A 1

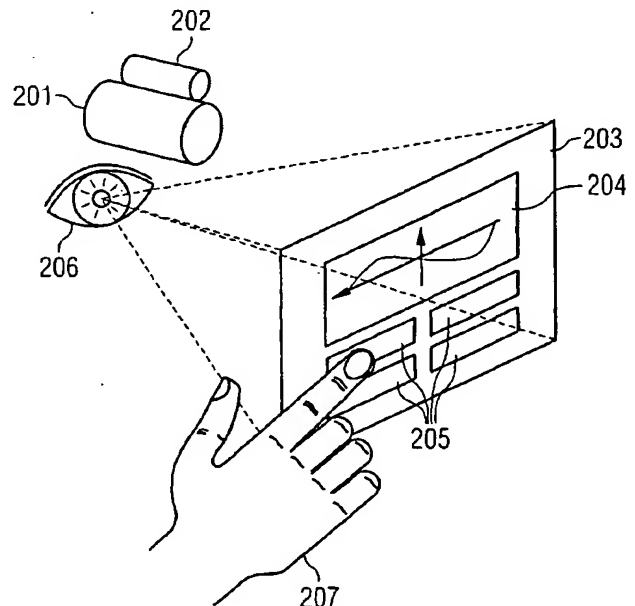
⑦1 Anmelder:  
Siemens AG, 80333 München, DE

⑦2 Erfinder:  
Feil, Henry, 86926 Greifenberg, DE

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

⑤4 Verfahren zur visuellen Darstellung und interaktiven Steuerung von virtuellen Objekten auf einem Ausgabe-Sichtfeld

⑤7 Die Erfindung betrifft ein Verfahren zur grafischen Visualisierung von Objekten, mit dessen Hilfe die auf mindestens einem Sichtfeld 203 dargestellten Objekt durch interaktive Steuerbefehle eines Benutzers in ihren Eigenschaften und/oder Aktionen auf eine komfortable und zuverlässige Weise manipuliert, gesteuert bzw. beeinflusst werden. Bei den dabei verwendeten Technologien zur Eingabe der vom Benutzer abgesetzten Steuerbefehle kann es sich neben den standardmäßigen manuell bedienbaren, mechanischen bzw. berührungssensitiven Eingabemechanismen über Tastatur, Maus, Trackball, Joystick, Grafiktablett und Griffel, taktile Displays etc. um Vorrichtungen 501, 502 und 503 zur Aufnahme, Erkennung, Interpretation und Verarbeitung akustischer und/oder optischer Signale eines Benutzers handeln. Der Benutzer ist somit nicht mehr auf das Vorhandensein zusätzlicher Hardware-Vorrichtungen zur manuellen Eingabe von Steuerbefehlen angewiesen. Die Auswertung der eingegebenen Information kann stattdessen bzw. zusätzlich mit Hilfe von Methoden der Signal- bzw. Mustererkennung erfolgen. Erst durch diesen erfinderischen Schritt wird aus einem reinen Ausgabe-Sichtfeld, bei dem keine Steuerungsmöglichkeit für den Benutzer vorgesehen ist, ein interaktiv bedienbares Ein- und Ausgabe-Sichtfeld 203. Die Art des Eingabeverfahrens kann dabei auf die individuell vorhandenen Fähigkeiten des Benutzers zugeschnitten sein.



DE 100 56 291 A 1

## A. Schilderung des allgemeinen Problems

[0001] Im Rahmen der Kommunikation zwischen Mensch und Maschine treten häufig Probleme auf, die auf der ungenügenden Anpassung der Maschine an die Eigenschaften der Aufnahme, Verarbeitung und Ausgabe von Information durch den menschlichen Anwender beruhen. Diese Fehlanpassung kann einerseits zu einer vom Anwender nicht mehr zu bewältigenden Informationsflut führen, vor allem wenn mehrere Aufgaben erledigt werden müssen. Andererseits kann auch eine Unterforderung des Anwenders, beispielsweise bei einer hochautomatisierten Anlage, bei der dem Menschen lediglich eine Kontrollfunktion zukommt, bewirken, dass durch die Monotonie der Arbeitssituation ein Leistungsabfall eintritt und Störfälle infolge mangelnder Übung der Vorgehensweise in solchen Situationen nicht mehr beherrscht werden. Auch die mangelnde Berücksichtigung der Kenntnisse und des Ausbildungszustands des Benutzers einer Maschine sind hier zu nennen. Das menschliche Verhalten wird, zum Beispiel bei der Auswahl, Bewertung und Verknüpfung von Information, bei der Entscheidungsfindung, beim Problemlösen sowie bei der Planung und Ausführung von Handlungen, nur unzureichend berücksichtigt und unterstützt, wenn es sich um die Auslegung technischer Systeme handelt.

[0002] Für die Anpassung technischer Systeme an den Menschen ist somit Vorwissen über seine Eigenschaften, seine Verhaltensmuster, seine Fertigkeiten und seinen Kenntnisstand notwendig. Im Zusammenhang mit der Mensch-Maschine-Kommunikation interessieren also seine sensorischen, kognitiven und motorischen Eigenschaften. Auf der Seite der durch die Sinneskanäle vorgegebenen sensorischen Eigenschaften des Menschen werden von herkömmlichen Maschinen und Geräten zur Ausgabe von Information im Wesentlichen die folgenden Kanäle angesprochen:

- der visuelle Kanal (Augen) durch optische Signale,
- der auditive Kanal (Ohren) durch akustische Signale und
- der taktile Kanal (Tastsinn) durch haptische Signale.

[0003] Nach der Verarbeitung der Signale im Gehirn (Kognition) stehen auf der Seite der durch die Ausgabekanäle vorgegebenen motorischen Eigenschaften des Menschen im Wesentlichen folgende Kanäle zur Verfügung:

- die Arm-, Hand- und Finger- bzw. Bein- und Fußmotorik sowie Körper-, Kopf-, Augen- oder Mundbewegungen, also physikalische Bewegungen, Gebärden, Gestik und Mimik für mechanische bzw. optische Signale,
- die Sprachmotorik für akustische Signale.

[0004] Über diese Kanäle können Signale in ein Informationssystem eingegeben werden, um eine gewünschte Aktion des Systems auszulösen.

## B. Bekannte Lösung des allgemeinen Problems nach dem aktuellen Stand der Technik

[0005] Das Ziel der Entwicklung von geeigneten Schnittstellen zwischen Mensch und Maschine ist es, von den Eigenschaften der menschlichen Kommunikationskanäle und Fertigkeiten auszugehen, um Geräte, Interaktionstechniken und Schnittstellen herzustellen, die eine effektive wechs-

selseitige Kommunikation über diese Kanäle gewährleisten. Um dieses Ziel zu erreichen, sind sogenannte "Virtuelle Realitäten" (VR) besonders geeignet. Unter dem Begriff "Virtuelle Realität" (VR) versteht man die computerbasierte Erzeugung einer intuitiv wahrnehmbaren oder empfindbaren Szene, bestehend aus ihrer grafischen Darstellung und den Interaktionsmöglichkeiten für den Benutzer. Eine virtuelle Umgebung ermöglicht einem Benutzer Zugriff zu Informationen, die andernfalls am gegebenen Ort oder zur gegebenen Zeit nicht verfügbar wären. Sie stützt sich auf natürliche Aspekte der menschlichen Wahrnehmung, indem sie visuelle Information in drei räumlichen Dimensionen einsetzt. Diese Information kann beispielsweise gezielt verändert oder mit weiteren sensorischen Reizen angereichert werden. Wesentliche Voraussetzungen sind die Kontrolle der Perspektive in Echtzeit und die Möglichkeit der aktiven Einflussnahme auf die dargestellte Szene durch den Benutzer des Systems.

[0006] Bei der Navigation durch virtuelle Umgebungen kann der Benutzer die für ihn natürliche Art der Steuerung einsetzen. Dies können beispielsweise entsprechende Arm- oder Beinbewegungen, Positionieren des Kopfes bzw. der Augen, Drehen des Körpers oder Zulaufen auf ein Objekt beinhalten. Durch den Einsatz bereits vorhandener Fertigkeiten des Benutzers zur Steuerung kann die kognitive Belastung während der Interaktion zwischen Mensch und Maschine reduziert werden. Dadurch kann die Bandbreite der Kommunikation zwischen Mensch und Maschine erhöht und die Bedienbarkeit der Maschine verbessert werden. Während bei den herkömmlichen Formen der Mensch-Maschine-Kommunikation die Steuerung der Maschine kommandoorientiert erfolgt, müssen bei der Steuerung von Objekten in virtuellen Umgebungen keine spezifischen Kommandos neu erlernt und eingesetzt werden: Der Computer "beobachtet" den Benutzer passiv und reagiert aufgrund von dessen Augen-, Kopf- und/oder Handbewegungen etc. unter Echtzeitbedingungen in angemessener Weise.

[0007] Bei den handelsüblichen VR-Applikationen unterscheidet man Systeme, bei denen der Benutzer vollständig in die virtuelle Umgebung integriert ist (engl.: "Immersion") und Systeme, die nur ein "Fenster" zur virtuellen Realität darbieten. Neben den bekannten Formen der Mensch-Maschine-Kommunikation wie

- Direktmanipulation von Objekten durch manuelle feinmotorische Operationen (Zeigen, Berühren, Greifen, Bewegen, Festhalten etc.),
- formale Interaktionssprachen (Programmiersprachen, Kommandosprachen und formale Abfragesprachen),
- natürlichsprachliche Interaktion,
- gestische Interaktion mittels nonverbaler symbolischer Kommandos (Mimik, Gestik, Gebärden, Bewegungen) sowie
- hybride aufgabenorientierte Interaktionsformen

kann man virtuelle Realitäten auch als eine neue Form der Mensch-Maschine-Kommunikation auffassen. Wie der Name "Virtuelle Realität" schon andeutet, ist hierfür eine gewisse Realitätstreue der Darstellung notwendig: Dem Anwender soll diejenige sensorische Information dargeboten werden, die zur Bearbeitung einer Aufgabe oder zum Erreichen eines Ziels erforderlich ist.

[0008] Die visuelle Wahrnehmung liefert nicht nur Informationen über die Lage, Bewegung, Form, Struktur, Kontur, Textur, Farbe bzw. Musterung von Objekten etc., sondern auch Informationen über die relative Körperstellung des Betrachters und dessen Bewegungen sowie über die Beschaf-

fenheit der dreidimensionalen Umgebung. Synthetisch generierte Umgebungen können dabei realistischer gestaltet werden, wenn möglichst viele der in natürlichen Umgebungen auftretenden Informationen (Bewegungsparrallaxe, Fluchtpunkte der perspektivischen Darstellung, räumliche Tiefenwirkung und Plastizität, Beleuchtung und Schattenwurf, Verdeckung, Glanzwirkung, Spiegelungseffekte und diffuse Reflexion etc.) simuliert werden. Wie viele und welche Informationen präsentiert werden sollen, hängt von der jeweiligen Aufgabenstellung ab. Die Unterschiede zwischen realer und virtueller Welt bestimmen, als wie realistisch die Simulation wahrgenommen wird.

[0009] Die visuelle Information muss zur Realisierung virtueller Realitäten durch einen Computer simuliert werden. Dabei sind ähnliche Aspekte relevant wie in der Malerei. Bei der computergestützten Simulation dreidimensionaler Welten wird üblicherweise die Projektion einzelner Lichtstrahlen simuliert. Ausgangspunkt einer solchen Simulation ist die Spezifikation der zu simulierenden Umgebung. Dazu müssen die einzelnen Objekte mit ihren Eigenschaften und ihrer Lage festgelegt werden. Zur Visualisierung werden dann die Intensitäten einzelner Bildpunkte berechnet und auf das Ausgabemedium projiziert.

[0010] Mit Hilfe dieser Simulationen lassen sich völlig neue Arten des Lernens und Übens realisieren (Beispiele: Fahrzeug- bzw. Flugzeugsimulator), andererseits wird dabei immer auch von bestimmten Aspekten der realen Welt abstrahiert. VR-Applikationen bewirken daher gleichzeitig eine Anreicherung und eine Einschränkung der Erfahrungsmöglichkeiten des Anwenders.

[0011] Grundsätzlich bestehen VR-Systeme aus Sensoren und Aktoren sowie deren Kopplung. Wichtige Hardware-Bestandteile sind unter anderem folgende:

- "Displays" zur Präsentation der virtuellen Umgebung. Im Rahmen der visuellen Präsentation kommen heute vor allem Monitore, "Head Mounted Displays" (HMD), "Binocular Omni-Oriented Monitors" (BOOM) und Projektionssysteme; verwendet werden aber auch auditive bzw. taktile Displays, welche auf akustische bzw. manuelle Benutzereingaben reagieren.
- Positionierungs- und Orientierungssysteme zur Erfassung von Standort und Perspektive des Benutzers. Hierbei wird unterschieden zwischen der Bestimmung der absoluten Position (engl.: "Position Tracking") und der Messung der Beugung von Gelenken (engl.: "Angle Measurement"). Zum Einsatz kommen elektromagnetische, kinematische, akustische, optische sowie bildverarbeitende Prozeduren.
- Interaktions- und Manipulationssysteme zum Agieren und Reagieren des Anwenders in der virtuellen Umgebung. Hierfür werden Zeigegeräte (2D- bzw. 3D-Mäuse, -Trackballs, -Joysticks etc.) bzw. taktile Geräte (Touchscreen, elektromagnetisches Grafiktablett und Griffel etc.) verwendet; auch sogenannte "Datenhandschuhe" mit Beugungs- und Drucksensoren werden in zunehmendem Maße eingesetzt. Auch Sprachsteuerung ist in diesem Zusammenhang zu nennen.
- Berechnungssysteme und Software zur Erzeugung der virtuellen Umgebung unter Echtzeitanforderungen.
- Netzwerke zur Integration verschiedener Benutzer, durch die sich neue Formen der Zusammenarbeit entwickeln können.

[0012] Die verschiedenartigen technischen Varianten von helm- bzw. kopfbasierten Systemen zur Visualisierung virtueller Realitäten werden im englischen Sprachgebrauch zusammenfassend als "Visually Coupled Systems" (VCS) be-

zeichnet. Sie bestehen aus den folgenden wichtigen Komponenten:

1. einem am Kopf bzw. Helm befestigtes Display,
2. einem Gerät zum Bestimmen der Kopf- und/oder Blickbewegungen des Anwenders,
3. einer Quelle visueller Information, die von der Kopf- und/oder Blickrichtung des Anwenders abhängt.

[0013] Beim Einsatz eines derartigen Systems für VR-Applikationen können sowohl Informationen aus der realen als auch aus der virtuellen Umgebung gleichzeitig dargeboten werden. Man spricht dabei von "See-Through Displays" für die Darstellung von angereicherten Realitäten.

[0014] Zur Visualisierung sind geeignete optische Komponenten (Linsen, halbdurchlässige Spiegel) hoher Qualität erforderlich, die ein scharf fokussiertes, vergrößertes Abbild der Bildquelle ermöglichen. Als Bildquellen kommen verschiedene Systeme in Betracht; am häufigsten werden jedoch Kathodenstrahlröhren oder LCD-Bildschirme eingesetzt. Wünschenswert sind dabei eine hohe Auflösung und Leuchtdichte, hohe Farbsättigung und hoher Kontrast sowie geringe Abmessungen der Bildquelle. Zur Visualisierung dreidimensionaler Objekte sind zwei derartige Bildquellen erforderlich.

[0015] Die Verfolgung von Kopfbewegungen ist ein wichtiger Bestandteil von VR-Applikationen. Üblicherweise werden Position und Orientierung des Kopfes im Raum ermittelt, fortgeschrittene Systeme können außerdem die Blickrichtung verfolgen. Die meisten Systeme setzen dazu entweder Ultraschall, magnetische oder Lichtenergie zur Kommunikation zwischen den am Kopf angebrachten Sendern und den Empfängern ein. Wichtige technische Daten, die bei der Auswahl dieser Systeme eine Rolle spielen, sind:

- die Anzahl der Freiheitsgrade für die Bewegungsrichtungen, welche registriert und verfolgt werden können,
- der erfassbare Winkelbereich,
- die statische Genauigkeit (Erschütterungsempfindlichkeit),
- das Auflösungsvermögen,
- die Zuverlässigkeit,
- der Datendurchsatz und die Bildschirm-Abtastrfrequenz,
- die Schnittstelle zum Computer sowie
- weitere Performanzaspekte.

[0016] VR-Applikationen lassen sich in einer Reihe von unterschiedlichen Bereichen in der Praxis erfolgreich anwenden. Im Folgenden werden einige Anwendungsmöglichkeiten beispielhaft skizziert.

- Entwicklung virtueller Prototypen, zum Beispiel in der Automobilindustrie: Wegen der zunehmenden Komplexität der zu entwickelnden Bauteile und immer kürzer werdenden Entwicklungszyklen ist es sinnvoll, Planung und Design mit Unterstützung von Computern zu betreiben. Es lassen sich eine höhere Innovationsrate, kürzere Entwicklungszyklen und eine bessere Koordination des in einem Team vorhandenen Wissens erreichen, wenn mehrere Personen parallel das Produkt in einer virtuellen Umgebung planen und designen. Durch den Einsatz virtueller Realitäten lassen sich auch komplexe dreidimensionale Gebilde leicht spezifizieren. Die Brauchbarkeit virtueller Prototypen kann anschließend von verschiedenen Benutzern evaluiert werden.

– Einsatz im Ausbildungsbereich: Durch das Erlernen des Umgangs mit (virtuellen) Objekten, interaktive Demonstrationen, Visualisierung abstrakter Konzepte, virtuelles Training des Verhaltens in gefährlichen Situationen, virtuelle Erforschung entfernter Orte oder Epochen können Wissen vermittelt, kreative Fertigkeiten geschult und Verhaltensmuster trainiert werden.

– Fahr- und Flugtraining in entsprechenden Simulatoren:

Durch den Einsatz von Simulatoren kann das Verhalten insbesondere in Notsituationen geschult werden

[0017] Die heute verfügbaren Technologien zur Eingabe von Information in ein datenverarbeitendes System lassen sich nach den verwendeten Sensoren in vier Gruppen einteilen: mechanische (z. B. Tastaturen, Mäuse, Trackballs und Joysticks), elektrische (z. B. taktile Displays und Grafiktablets), optische (z. B. Lichtgriffel) und akustische (z. B. Spracheingabe- und Sprachinterpretationssysteme). Im Folgenden soll auf die nach dem heutigen Stand der Technik gebräuchlichen Hilfsmittel zur Eingabe von Information, die zur Steuerung von Objekten im Bereich von VR-Applikationen eingesetzt werden, kurz eingegangen werden.

[0018] Sowohl Tastaturen, Mäuse, Trackballs als auch Joysticks benötigen eine Ablagefläche, also einen festen Standplatz. Mit einem Touchscreen dagegen ist es möglich, direkt auf Objekte, die auf dem Bildschirm abgebildet sind, mit dem Finger zu zeigen, ohne weitere platzraubende Zusatzgeräte auf dem Schreibtisch zu benötigen. Niedrig auflösende Touchscreens weisen 10 bis 50 Positionen in waagerechter und senkrechter Richtung auf und benutzen eine horizontale und vertikale Reihe von Infrarot-Leuchtdioden und Fotosensoren, um ein Gitter von unsichtbaren Lichtstrahlen unmittelbar vor dem Bildschirm aufzubauen. Bei einer Berührung des Bildschirms werden sowohl vertikale als auch horizontale Lichtstrahlen unterbrochen. Aus dieser Information kann die aktuelle Fingerposition ermittelt werden.

[0019] Eine andere bekannte Ausführungsform berührungssensitiver Informationseingabegeräte ist das kapazitiv gekoppelte Touch-Panel. Dieses liefert eine Auflösung von ca. 100 Positionen in jeder Richtung. Wenn ein Benutzer die leitfähig beschichtete Glasplatte des Touchscreens mit einem Finger berührt, kann aufgrund der Impedanzänderung die aktuelle Fingerposition ermittelt werden. Andere hochauflösende Panels verwenden zwei minimal voneinander entfernte, transparente Schichten. Eine davon ist leitfähig beschichtet, die andere mit einem Widerstandsmaterial beschichtet. Durch den Anpressdruck des Fingers berühren sich diese beiden Lagen, und durch Messung des daraus resultierenden Spannungsabfalls kann dann die aktuelle Fingerposition ermittelt werden. Eine niedriger auflösende und billigere Variante dieser Technologie verwendet anstelle dieser Schichten ein Gitter von feinen Drähten.

#### C. Unzulänglichkeiten, Auswirkungen und Nachteile der bekannten Lösung

[0020] Die derzeit verwendeten Systeme zur Darstellung und Steuerung von Objekten in virtuellen Umgebungen berücksichtigen zwar in zunehmendem Maße die Fähigkeiten des Menschen zur Aufnahme und Verarbeitung von Information, jedoch weisen sie einen entscheidenden Nachteil auf: Bei der Eingabe von Steuerbefehlen zur direkten Beeinflussung der dargestellten Szene ist der Benutzer nach wie vor auf herkömmliche Methoden zur manuellen Eingabe von Information angewiesen, wie zum Beispiel über Maus, Trackball, Joystick, Grafiktablett mit Griffel, Touchscreen. Die hierbei notwendigen Eingabemechanismen müssen

vom Benutzer erst erlernt werden, um auch in einer angemessenen Reaktionsgeschwindigkeit ausgeführt werden zu können. Dagegen werden die angeborenen bzw. bereits vorhandenen erlernten Fähigkeiten des Menschen zur Kommunikation mittels akustischen Signalen (z. B. Sprache) bzw. optischen Signalen (z. B. Mimik, Gestik, Gebärden und Bewegungen) bei der Eingabe von Information zur Steuerung von Objekten nur unzureichend berücksichtigt.

[0021] Die Manipulation der Eigenschaften und Beeinflussung der Aktionen von Objekten einer dargestellten Szene setzt ein kompliziertes Zusammenspiel von Sensorik, kognitiver Verarbeitung und Motorik voraus, auf das viele Faktoren einwirken (individuelle Verhaltensmuster und Fähigkeiten, Erfahrungen, Umwelteinflüsse etc.). Bei Interaktionen in einer virtuellen Welt kommen noch zusätzliche Schwierigkeiten hinzu. Zur Steuerung, Manipulation bzw. Beeinflussung von Objekten ist insbesondere eine reflexartige bzw. kognitive sensorisch-motorische Rückkopplung wichtig, die beispielsweise von Rezeptoren in der Haut, kinästhetischen Empfindungen, dem Gleichgewichtssinn sowie visuellen und/oder akustischen Empfindungen stammt. Dabei ergibt sich in vielen Fällen eine notwendige Redundanz, die bei VR-Applikationen nicht immer gegeben ist. Aufgrund der unzureichenden sensorischen Rückkopplung bei VR-Applikationen wird zudem das Erlernen motorischer Fertigkeiten erschwert.

#### D. Durch die Erfindung zu lösende spezielle Aufgabe

[0022] Ein ideales Medium zur Kommunikation zwischen einem Benutzer und einem Informationssystem sollte sowohl auf die sensorischen und perzeptuellen als auch auf die motorischen Fähigkeiten sowie auf die spezifischen Eigenschaften des menschlichen Benutzers abgestimmt sein. Dabei sollte die Information so strukturiert sein, dass eine optimale Übereinstimmung zwischen der Repräsentation der ausgegebenen Information und dem mentalen Modell des Benutzers erzielt wird: Werden die dem Benutzer anzuzeigenden Informationen in einer Weise dargeboten, dass zum Beispiel sein räumliches Wahrnehmungsvermögen angesprochen wird, kann der Benutzer mit erstaunlich komplexen Informationsmengen pro Zeiteinheit umgehen. Ebenso sollte das Informationssystem in der Lage sein, möglichst viele Arten der von einem Benutzer abgesetzten Information aufzunehmen, zu verstehen und zu verarbeiten und in entsprechende Aktionen umzusetzen. Damit ist der Vorteil verbunden, dass der Benutzer effizienter und schneller auf neue Ereignisse und Situationen reagieren kann. Benutzerfreundlichkeit und Aufgabenangemessenheit sind somit typische Merkmale, über die ein ideales Medium verfügt. Diese Merkmale können sich wie folgt äußern:

- Übereinstimmung zwischen Art, Umfang und Ausgabegeschwindigkeit und Präsentation der ausgegebenen Information mit den sensorischen Eigenschaften des menschlichen Benutzers,
- Berücksichtigung aller Informationskanäle des Benutzers bei der Aufnahme, Erkennung und Interpretation empfangener Steuersignale des Benutzers,
- leichte Erlernbarkeit und intuitive Bedienbarkeit des Mediums,
- hohe Bandbreite der Informationsübermittlung zum Gehirn und hoher Durchsatz der Information,
- dynamische Adaption der Applikation an die individuellen Eigenschaften, Fähigkeiten, Aufgaben, Arbeits- und Organisationstechniken des Benutzers,
- Verwendung einer natürlichen Interaktionssprache mit hohem semantischem Inhalt,



- Zuverlässigkeit, Robustheit und Wartbarkeit des Mediums,
- soziale Akzeptanz des Mediums in der Bevölkerung,
- Berücksichtigung gesundheitlicher, ergonomischer und sicherheitsrelevanter Aspekte etc.

[0023] Der vorgelegten Erfindung liegt die Aufgabe zugrunde, die bestehende Situation mit technischen Mitteln zu verbessern. Daher widmet sich die Erfindung vorrangig der Aufgabe, komfortable und zuverlässig arbeitende Verfahren herbeizustellen, mit deren Hilfe dem Anwender eine aktive Steuerung virtueller Objekte ermöglicht wird, wobei die bereits vorhandenen Fertigkeiten des Benutzers zur Ausnutzung von Information genutzt werden. Diese Aufgabe wird durch ein Verfahren mit Merkmalen nach Patentanspruch 1 gelöst.

E. Erfindungsgemäße Lösung der speziellen Aufgabe gemäß den Patentansprüchen und Vorteile dieser Lösung

#### Zu Patentanspruch 1

[0024] Der Benutzer soll in die Lage versetzt werden können, mit Hilfe geeigneter Interaktionsmöglichkeiten die Eigenschaften und/oder Aktionen der dargestellten Objekte im Sichtfeld 203 zu manipulieren bzw. zu beeinflussen. Zu diesem Zweck sind zusätzliche hardwaretechnische Hilfsmittel zur manuellen Eingabe von Steuerbefehlen (z. B. via Tastatur, Maus, Trackball, Joystick oder Touchscreen) nicht unbedingt erforderlich. Die vorliegende Erfindung greift deshalb zur Erfassung und Verarbeitung vom Benutzer in Form von Signalen ausgesendeter Information auf Methoden der Signal- bzw. Mustererkennung zurück. Erst durch diesen erfinderischen Schritt wird aus einem reinen Ausgabe-Sichtfeld ein interaktiv bedienbares Ein- und Ausgabe-Sichtfeld 203, das an die Fähigkeiten des Benutzers angepasst ist. Bei den durch den Benutzer manipulierbaren Eigenschaften von Objekten einer dargestellten Szene 204 kann es sich beispielsweise um Merkmale wie Ortsvektor, Lage, Ausmaße, Betrachtungsperspektive, geometrische Form, Struktur, Kontur bzw. Textur der Objekte handeln. Vorstellbar sind aber auch Interaktionsmöglichkeiten, mit denen die Art der Darstellung der Objekte verändert wird, wie beispielsweise Farbe bzw. Musterung, Helligkeit, Kontrastwirkung gegenüber dem Hintergrund, Glanzeffekte, Spiegelungen und Reflexionen, Schattierung etc. Bei den durch den Benutzer beeinflussbaren Aktionen von Objekten der dargestellten Szene kann es sich etwa um die Bewegung der Objekte handeln. Die Bewegung der Objekte kann dabei Translations- und/oder Rotationsanteile beinhalten, wobei die kinematischen Daten der Objekte, wie zum Beispiel Betrag und/oder Richtung seines Geschwindigkeits- bzw. Beschleunigungsvektors bzw. seines Winkelgeschwindigkeits- bzw. Winkelbeschleunigungsvektors, individuell gemäß den Steuerbefehlen des Benutzers verändert werden können.

#### Zu Patentanspruch 2

[0025] Die Registrierung der interaktiven Steuerbefehle des Benutzers kann dabei mit Hilfe von Sensoren 502 und/oder Aufnahmegeräten 501 erfolgen. Nach der Datenerfassung können die erfassten Eingabe-Daten einer Auswertung und Steuerungsvorrichtung 503 zugeführt werden, als Steuerbefehle interpretiert und verarbeitet werden. Um die Möglichkeit einer Fehlinterpretation von ausgesendeten Signalen eines Benutzers durch das Datenverarbeitungssystem möglichst gering zu halten und damit die Zuverlässigkeit des Systems zu erhöhen, muss das System an die gewohnten

Eigenschaften des Benutzers angepasst werden. Aus diesem Grund müssen dem System alle möglichen Eingangssignale eines individuellen Befehlssatzes des betreffenden Benutzers in einer Trainingsphase mehrfach wiederholt "beigebracht" werden. Entsprechend dieser vom System als Steuerbefehle interpretierten Signale ist es möglich, die Objekte einer dargestellten Szene zu manipulieren bzw. zu beeinflussen oder/und eine Aktion auszulösen. Die auf diese Weise manipulierten bzw. beeinflussten Objekte oder/und die damit ausgelösten Aktionen können auf einem Sichtfeld 203 dem Benutzer grafisch visualisiert bzw. akustisch und/oder optisch angezeigt werden. Dadurch ist ein Rückkopplungskreis zwischen Ausgabe- und Eingabedaten gegeben, mit dessen Hilfe die Reaktion auf Änderungen von Objekten einer dargestellten Szene 204 nahezu in Echtzeit ermöglicht werden kann.

#### Zu Patentanspruch 3

[0026] Bei den im Rahmen dieses Verfahrens steuerbaren Objekten kann es sich beispielsweise um real existierende Objekte in einer realen Umgebung handeln. Ein typisches Anwendungsbeispiel ist etwa die Fernsteuerung eines Roboters zur Durchführung gefährlicher Arbeiten an schwer zugänglichen und/oder entfernten Orten, zum Beispiel zur Friedigung von Reparaturarbeiten in Kanalisationsrohren, zur Untersuchung radioaktiver Substanzen in hermetisch abgeschotteten Hochsicherheitsräumen oder zur Datengewinnung und -übermittlung bei unbemannten Weltraummissionen. Ein anderes typisches Anwendungsbeispiel ist die Möglichkeit der Fernsteuerung von Arbeits- und/oder Haushaltsgeräten mittels nicht-manueller Eingabeverfahren für körper- oder/und manuell behinderte Menschen. Die Art des Eingabeverfahrens kann dabei auf die individuell vorhandenen Fähigkeiten des Benutzers zugeschnitten sein.

#### Zu Patentanspruch 4

[0027] Vorstellbar sind aber auch Ausführungsbeispiele dieser Erfindung, bei denen die zu steuernden Objekte nicht real existierende Objekte in einer virtuellen Umgebung eines computergesteuerten Modells sind. Vor allem im Ausbildungsbereich erscheint der Einsatz von Systemen zur interaktiven Steuerung virtueller Objekte sinnvoll: Durch das Erlernen des Umgangs mit virtuellen Objekten, interaktive Demonstrationen, Visualisierung abstrakter Konzepte, virtuelles Training des Verhaltens in gefährlichen Situationen, virtuelle Erforschung entfernter Orte oder Epochen können Wissen vermittelt, kreative Fertigkeiten geschult und Verhaltensmuster trainiert werden. Typische Einsatzgebiete sind beispielsweise Flugsimulatoren, mit deren Hilfe kritische Situationen im Flugverkehr (Triebwerksausfälle, Strömungsabriss, Notlandungen auf dem Wasser und auf dem Land etc.) von auszubildenden Piloten trainiert werden und Lernfortschritte quantitativ beurteilt werden können.

#### Zu Patentanspruch 5

[0028] Eine natürliche Form der Interaktion von Mensch und Maschine sowie der Eingabe von Information in ein informationsverarbeitendes System ist die der natürlichsprachlichen Kommunikation. Aus diesem Grund bietet sich diese Form der Eingabe für Ausführungsbeispiele der vorliegenden Erfindung an. Die bisher und in absehbarer Zeit verwendeten Spracheingabesysteme machen jedoch nur in einem eingeschränkten Umfang von der natürlichsprachlichen Kommunikation Gebrauch, indem nur Wörter eines vom Hersteller fest vorgegeben Basisvokabulars geringen

Umfangs zuzüglich eines benutzerspezifischen Fachvokabulars (in der Regel einige hundert bis tausend Wörter) oder auch Zusammensetzungen dieser Wörter zugelassen werden. Zur Erhöhung der Erkennungsrate muss ein Anwender sowohl das Basisvokabular als auch sein benutzerspezifisches Fachvokabular dem System in einer Trainingsphase mehrfach wiederholt vorsprechen, damit sich das System auf die Stimme des Benutzers und etwaige Abweichungen vom Normalzustand der Stimme, zum Beispiel infolge Erkältung oder Heiserkeit des Benutzers, einstellen kann. Bei einer Ausführungsform spracherkennender Systeme werden während dieser Trainingsphase vom System statistische Modelle (in der Regel "Hidden-Markov-Modelle") generiert, mit deren Hilfe beispielsweise der Abfolge einzelner Phoneme bzw. Silben der trainierten Wörter Wahrscheinlichkeitsdichtefunktionen zugeordnet werden, so dass neu vorgesprochene Wörter mit einer vorhersagbaren Wahrscheinlichkeit als richtig bzw. falsch erkannt werden.

#### Zu Patentanspruch 6

[0029] Bei der Vorrichtung 503 zur Erkennung und Interpretation der Steuerbefehle des Benutzers kann es sich beispielsweise um ein automatisches Erkennungssystem für akustische Signale handeln, speziell für Sprachsignale. Die meisten der heute eingesetzten automatischen Spracherkennungssysteme sind zum Beispiel wort-, silben- oder phonembasierte Systeme, die nur für kontextunabhängige Spracherkennung konzipiert sind. Das bedeutet, dass die aus Wörtern eines trainierten Vokabulars bestehenden Sprachkommandos dem System derzeit mit ausreichend großen Sprechpausen ("diskret") vorgesprochen werden müssen, um für den Benutzer akzeptable Worterkennungsraten im Bereich von ca. 90% bis 95% zu erzielen. Handelsübliche polyphonbasierte Prototypen kontextabhängiger Spracherkennungssysteme für fließend ("kontinuierlich") gesprochene Sprache erreichen heute schon Worterkennungsraten von ca. 95% bis 98%. Allerdings ist die Rechenleistung des Systems für kontinuierliche kontextabhängige Spracherkennung so groß, dass sie auch auf den derzeitigen Hochleistungsrechnern nicht mehr in Echtzeit durchgeführt werden kann.

#### Zu Patentanspruch 7

[0030] Eine weitere Form der Mensch-Maschine-Kommunikation, welche den Eigenschaften und Fertigkeiten des Benutzers entgegenkommt, ist die der Auswertung von optischen Signalen, wie etwa Körper-, Kopf-, Gesichts-, Bein-, Fuß-, Arm-, Hand- und/oder Fingerbewegungen, also Gebärden, Gestik und/oder Mimik. Die weitverbreiteten Fähigkeiten des Menschen, Informationen durch die Position, Raumrichtung und Bewegung von Körperteilen zu kodieren, stellt heute noch eine große Herausforderung für die Gestaltung von Eingabeverfahren für die Mensch-Maschine-Kommunikation dar.

#### Zu Patentanspruch 8

[0031] Zum Empfang von optischen Signalen des Benutzers können Sensoren 202 und/oder Aufnahmegeräte 201 verwendet werden. Konkret kann es sich dabei um optische bzw. akustische Abstandssensoren und/oder Videokameras handeln. Der Abstandssensor kann über eine Ultraschall- bzw. hochfrequente Strahlungsquelle verfügen sowie einem Detektor und Vorrichtungen, die den Schall bzw. die Strahlung auf das Messobjekt bündeln und die vom Messobjekt reflektierten Schall- bzw. Lichtwellen auf dem Detektor

sammeln. Dabei kann ein Teil des Umgebungsschalls bzw. der Umgebungsstrahlung mit Hilfe von Filtern abgeblockt werden. Damit das Sensorsystem unabhängig von den Umgebungsbedingungen zuverlässig funktioniert, kann eine Kompression, Kodierung und Modulation der von der Signalquelle ausgesendeten Signale und eine geeignete Signalverarbeitung auf der Detektorseite zur Demodulation, Dekodierung und Dekompression der empfangenen Signale vorgesehen sein.

#### Zu Patentanspruch 9

[0032] Bei der Vorrichtung 503 zur Interpretation von interaktiven Steuerbefehlen eines Benutzers kann es sich beispielsweise auch um ein automatisches Erkennungssystem für Bewegungssignale handeln, wobei Methoden der Bildverarbeitung zum Einsatz gelangen können. Eine Möglichkeit ist zum Beispiel, die Lippenbewegungen des Benutzers beim Sprechen zu analysieren, um damit die Erkennungssicherheit der Spracheingabe zu erhöhen. Moderne Untersuchungen haben gezeigt, dass durch die Kombination von akustischer und optischer Spracherkennung im Vergleich zur ausschließlich akustischen Spracherkennung die Wortfehlerrate bei einzelnen Sprechern um 30% bis 50% gesenkt werden kann. Andere Ausführungsformen können die Kopf- und/oder Augenbewegungen zur Informationseingabe verwenden. Hierzu ist eine Vermessung der Kopf- und/oder Blickrichtung erforderlich. Bei einem seit 1982 bekannten Eingabeverfahren, das auf einer Messung der Blickrichtung des Anwenders durch berührungslose Erfassung der Kopfbewegung beruht, erfolgt mittels eines vor einem Auge angebrachten Visiers die visuelle Rückkopplung der Kopfposition. Weiter ist am Kopf eine kleine Lichtquelle angebracht, deren Position mit Hilfe einer Videokamera vermessen wird, so dass Kopfbewegungen um zwei Drehachsen (horizontal und vertikal) erfasst werden können. Ein anderes Eingabeverfahren, das auf der Messung der Augenbewegungen basiert, ist seit 1987 bekannt. Dabei wird davon ausgegangen, dass die Auswahl eines auf einer optischen Anzeige dargestellten Objekts eine der häufigsten Eingabeoperationen ist und dass die visuelle Fixierung eines auszuwählenden Objekts eine normale Verhaltensweise des Menschen ist. Diese Art der Eingabe wird vor allem bei hohen Anforderungen an die Geschwindigkeit der Eingabe bzw. bei Behinderung der Hände oder Belegung durch andere Aufgaben empfohlen.

#### Zu Patentanspruch 10

[0033] Bei der Ausgabevorrichtung 505 zur Darstellung einer virtuellen Umgebung kann es sich beispielsweise um ein Gerät zur Erzeugung eines sogenannten "Virtual Retinal Display" (VRD) handeln, bei dem eine virtuelle Bildfläche 203 auf die Netzhaut 206 des Anwenders projiziert wird. Anstelle eines Bildschirms oder einer Anzeigevorrichtung wird lediglich eine kohärente Photonenstrahlung aussendende Lichtquelle benötigt. Dabei kann es sich um ein Gerät aus dem Arbeits- oder Freizeitbereich bzw. um eine Orthese handeln, die eine Erweiterung des menschlichen Sehvermögens bewerkstelligt. Gegenüber realen Displays weist ein VRD folgende Vorteile auf: 1. Die Auflösung des VRD wird lediglich durch die Beugung und optische Aberration des Lichtstrahls im menschlichen Auge begrenzt, nicht aber durch die bei Bildschirmen oder realen Displays technisch realisierbare Größe eines elementaren Bildelements (Pixelgröße). Aus diesem Grund können Bilder mit einer sehr hohen Auflösung generiert werden. 2. Die mit Hilfe eines VRDs erzielbare Bildhelligkeit kann durch die Intensität des ausgesendeten Lichtstrahls gesteuert werden. Bei Verwen-

dung eines Lasers als Lichtquelle kann die Bildhelligkeit des VRDs groß genug eingestellt werden, um im Freien eingesetzt werden zu können. 3. VRDs können entweder in einem Modus für virtuelle Realitäten oder in einem Modus für angereicherte Realitäten ("See-Through Modus") betrieben werden. 4. Die industrielle Fertigung einer Lichtquelle zur Generierung eines VRDs ist verhältnismäßig einfach und im Vergleich mit konventionellen Bildschirmen und Anzeigevorrichtungen mit geringen Produktionskosten durchführbar. 5. Da ein Großteil des generierten Lichts auf die Netzhaut des Betrachters fokussiert wird, arbeiten VRDs mit einem hohen Wirkungsgrad und weisen einen im Vergleich zu Bildschirmen und realen Displays geringen Leistungsverbrauch auf.

#### Zu Patentanspruch 11

[0034] Die virtuelle Bildfläche 401 kann beispielsweise virtuelle eingabesensitive Referenzpunkte 404, 405, 406 und 407 und/oder -flächen 403 in einer virtuellen Bildebene vorgegebener Raumrichtung aufweisen. Dies können ebene Flächen mit vorgegebenem Flächeninhalt sein. Die Raumrichtung der virtuellen Bildebene kann vorzugsweise senkrecht zur Blickrichtung orientiert sein, also dergestalt, dass der Normalenvektor  $n$  der virtuellen Bildebene parallel bzw. antiparallel zum Blickrichtungsvektor  $b$  liegt. Eingaben werden jeweils dann ausgelöst, wenn ein reales Objekt (ein Gegenstand oder ein Körperteil des Benutzers) die Ebene des virtuellen Displays an einer solchen eingabesensitiven Fläche 205 berührt bzw. die Fläche in einem beliebigen Eintrittswinkel schneidet. Mit Hilfe von Methoden der Mustererkennung kann es ermöglicht werden, eine beabsichtigte Eingabe von einer unbeabsichtigten Eingabe zu unterscheiden. Dabei könnte beispielsweise der Zeigefinger der rechten Hand eines Benutzers als einziges Eingabemodum erkannt werden. Eine Interaktion wird in diesem Fall nur dann zugelassen, wenn dieser Finger eine eingabesensitive Fläche in einer virtuellen Bildebene schneidet bzw. berührt.

#### Zu Patentanspruch 12

[0035] Der Abstand des Lotfußpunkts der virtuellen Bildebene zum Punkt schärfster Abbildung auf der Netzhaut eines der Augen 304 des Betrachters bzw. die Begrenzung und räumliche Ausrichtung der Bildfläche 303 der virtuellen Bildebene kann durch die Erkennung von optischen Signalen des Benutzers mit Hilfe eines optischen oder akustischen Abstandssensors 502, einer Videokamera 501 sowie einer nachgeschalteten Auswerteelektronik 503 vorgenommen werden.

#### Zu Patentanspruch 13

[0036] Bei den optischen Anzeigevorrichtungen wird neben konventionellen analogen und digitalen Displays zunehmend der Bildschirm eingesetzt, da hier die Kodierung der Information durch viele verschiedene Arten (Text, Graphik, Bild, Animation etc.) erfolgen kann. Außerdem stehen hier die Möglichkeiten der flexiblen Organisation der Information zur Verfügung, mit denen der erforderliche Such- und Auswahlvorgang einzelner Objekte unterstützt werden kann. Zur Darstellung räumlicher Sachverhalte können 3D-Anzeigen bzw. 3D-Brillen verwendet werden. Bei allen diesen optischen Anzeigen sind bei der Ablesung neben der dazu nötigen Aufmerksamkeit des Benutzers Bewegungen des Kopfes und/oder der Augen erforderlich. Zur Verhinderung der visuellen Ablenkung des Benutzers können, zum Beispiel bei der Flugzeug- oder Fahrzeugführung, Projekti-

onsvorrichtungen ("Head up"-Anzeigen) verwendet werden, bei denen die abzulesende Information in den Sichtbereich eingeblendet wird, auf den das Auge zur Wahrnehmung der Hauptaufgabe (etwa der Flugzeug- oder Fahrzeugführung) vorzugsweise ausgerichtet ist. Eine Sonderform ist die Ausführung als helmbasierte Anzeigevorrichtung in Form eines "Head-Mounted Displays" (HMD), bei der die Anzeigeeinheit im Helm des Benutzers befestigt ist. Herkömmliche VR-Anwendungen verwenden HMDs zur Darstellung virtueller Welten über ein Stereodisplaysystem sowie über Raumklangsysteme oder an das HMD gekoppelte Kopfhörer. Damit die dreidimensionale virtuelle Umgebung dem Betrachter auch bei Kopfbewegungen stabil erscheint, muß die Kopfbewegung in den sechs Freiheitsgraden erfaßt und dem Display weitergegeben werden. Dies muß mit einer genügend hohen Frequenz von mindestens 70 Hz erfolgen. Zur Messung der Kopfbewegung werden bisher eine Vielzahl verschiedener Messsysteme eingesetzt. Viele sind jedoch entweder für den Betrachter unbequem oder in der Beschaffung zu teuer. Zusätzliche können akustische bzw. haptische Ausgabevorrichtungen vorhanden sein, da akustische und haptische Signale geringer Intensität keine visuelle Ablenkung des Benutzers bewirken. Allerdings sind haptische Signalgeber derzeit nur als Blindenhilfsmittel in breitem Einsatz.

#### F. Figurenbeschreibung

[0037] Im Folgenden wird die Erfindung anhand bevorzugter Ausführungsbeispiele, wie sie in den Fig. 1 bis 5 geschildert sind, näher beschrieben.

[0038] Im Detail zeigen

[0039] Fig. 1 ein Designmuster 101, eine Prinzipdarstellung 102 sowie ein typisches Anwendungsbeispiel 103 eines

"Virtual Retinal Displays" (VRD) aus der Sicht des Benutzers (Quelle: "Technologies to Watch", Ausgabe Mai 2000),

[0040] Fig. 2 eine schematische Darstellung des Eingabeverfahrens über ein VRD,

[0041] Fig. 3 eine schematische Darstellung zur Erfassung des Abstands der virtuellen Bildebene bei einem VRD,

[0042] Fig. 4 ein Beispiel zur Erfassung der virtuellen Eckpunktkoordinaten 404, 405, 406 und 407 der Bildfläche 401 eines VRD und

[0043] Fig. 5 einen schematischen Hardwareaufbau zur Erfassung der Eingabe und Steuerung der Ausgabe eines VRD.

[0044] In Fig. 1 ist ein Designmuster 101 eines Microvision-Systems zur Erzeugung eines VRDs, eine Prinzipdarstellung 102 der Funktionsweise des VRDs sowie ein typisches Anwendungsbeispiel 103 eines VRDs dargestellt. Bei dem dargestellten Anwendungsbeispiel handelt es sich um eine Spannungsmessung im Kraftfahrzeugbereich, bei der dem Anwender über ein VRD Seriennummer, Profil, Messwerte und die gemessene Spannungs-Zeit-Kennlinie des geprüften Bauteils eingeblendet werden.

[0045] In Fig. 2 wird das Verfahren der Eingabe über berührungssensitive Schaltflächen 205 bei VRDs demonstriert. Abgebildet ist die Bildfläche 203 der virtuellen Bildebene mit den darin enthaltenen Teilbereichen für die virtuelle Ausgabefläche 204 und die virtuellen Taster bzw. Schalter 205. Werden die Eingabeflächen von einem Finger des Benutzers 207 oder einem Gegenstand berührt bzw. in einem beliebigen Eintrittswinkel durchstoßen, wird eine Eingabe erkannt und entsprechende Aktion ausgelöst. Die Eingaben des Benutzers können mit Hilfe einer Videokamera 201 und/oder einem Abstandssensor 202 erfasst werden.

[0046] Fig. 3 veranschaulicht bildhaft das Verfahren zur Bestimmung des Abstands der virtuellen Bildebene zur

Netzhaut eines der Augen 304 des Betrachters bei einem VRD. Durch die Position und räumliche Ausrichtung der ausgestreckten Handfläche 305 des Benutzers werden der Aufpunkt und die beiden Richtungsvektoren der virtuellen Bildebene 303 bestimmt. Zu diesem Zweck kann beispielsweise ein Ultraschall- bzw. Infrarotwellen aussendender Abstandssensor 302, eine Videokamera 301 sowie ein Verfahren zur Signal- bzw. Mustererkennung herangezogen werden. Idealerweise wird die räumliche Orientierung der virtuellen Bildebene vom Benutzer dergestalt festgelegt, dass der Normalenvektor  $n$  der virtuellen Bildebene parallel bzw. antiparallel zum Blickrichtungsvektor  $b$  verläuft, also das Skalarprodukt der betragsnormierten Vektoren  $n/||n||_2$  und  $b/||b||_2$  die Werte  $+1$  bzw.  $-1$  ergibt. Der Abstand der virtuellen Ebene zum Betrachter ergibt sich dann als der Abstand zwischen dem Lotfußpunkt, also dem Schnittpunkt zwischen virtueller Bildebene und Blickrichtungsgerade, und dem Punkt schärfster Abbildung auf der Netzhaut eines der Augen 304 des Betrachters.

[0047] In Fig. 4 wird gezeigt, wie die Begrenzungspunkte einer virtuellen polygonalen Bildfläche 401 eingegeben werden. In dem skizzierten Beispiel wurde über Vorgabe von vier Eckpunkten eine rechteckige Bildfläche definiert. Dabei können die kartesischen Koordinaten der Bildflächen-Eckpunkte 404, 405, 406 und 407 an denjenigen Raumpositionen generiert werden, an denen beispielsweise ein Finger einer Hand 408 des Benutzers oder ein Gegenstand die virtuelle Bildebene 401 berührt bzw. schneidet. Durch die Pfeile ist dabei die Bewegung des Fingers von einem Eckpunkt zum nächsten Eckpunkt der virtuellen Bildfläche angedeutet. Über die virtuellen Schalter bzw. Taster 403 können beispielsweise die Funktionen für das Ein- bzw. Ausschalten des auf der virtuellen Ausgabefläche 402 dargestellten Bildes, für das Umschalten zwischen mehreren Informationskanälen, für die Modus-Auswahl zur Einstellung von System- und/oder Bildparametern sowie für das Archiv zum Speichern von Ton- und/oder Bildsequenzen gesteuert werden.

[0048] Fig. 5 zeigt den Rückkopplungskreis zwischen Mensch und Maschine, der den schematischen Hardwareaufbau zur Erfassung der Eingabe und Steuerung der Ausgabe bei einem VRD enthält. Die über eine Videokamera 501 bzw. einen Abstandssensor 502 erhaltenen optischen Informationen des Benutzers werden dabei einer zentralen Steuerungseinheit 503 zugeführt, in der die Erkennung, Interpretation und Verarbeitung der Informationen erfolgt. Nach Auslösung einer entsprechenden Aktion werden die neu ermittelten bzw. berechneten Daten an die Steuerungseinheit 504 zur Steuerung des VRD weitergeleitet. Dieser übernimmt dann die Ausgabe der Text- und/oder Bilddaten über ein beispielsweise laserbetriebenes VRD 505. Nach Aufnahme und Verarbeitung dieser Ausgabedaten durch den menschlichen Benutzer 506 kann dann eine neue Reaktion des Benutzers auf den veränderten Ist-Zustand einsetzen.

[0049] Die Bedeutung der in den Fig. 1 bis 5 mit Ziffern bezeichneten Symbole kann der nachfolgenden Bezugszeichenliste entnommen werden.

#### Bezugszeichenliste

##### Nr. Bezeichnung

- 101 Designmuster eines Microvision-Systems zur Erzeugung eines "Virtual Retinal Displays" (VRD)
- 102 Prinzipdarstellung der Funktionsweise eines VRD-Systems
- 103 Anwendungsbeispiel eines VRD-Systems aus der Sicht des Anwenders
- 201 Videokamera

- 202 Optischer bzw. akustischer Abstandssensor
- 203 Bildfläche in der virtuellen Bildebene des VRD (virtuelles Sichtfeld bzw. virtuelles Display)
- 204 Virtuelle Ausgabefläche
- 205 Berührungssensitive Eingabeflächen (virtuelle Tasten bzw. virtuelle Schalter)
- 206 z. B. rechtes Auge des Benutzers
- 207 Auslösung einer Eingabe über das VRD durch Berührung bzw. Durchstoßung einer berührungssensitiven Eingabefläche durch einen Finger des Benutzers
- 301 Videokamera
- 302 Optischer bzw. akustischer Abstandssensor
- 303 Bildfläche in der virtuellen Bildebene des VRD (virtuelles Sichtfeld bzw. virtuelles Display)
- 304 z. B. rechtes Auge des Benutzers
- 305 z. B. rechte Hand des Benutzers
- 401 Bildfläche in der virtuellen Bildebene des VRD (virtuelles Sichtfeld bzw. virtuelles Display)
- 402 Virtuelle Ausgabefläche
- 403 Berührungssensitive Eingabeflächen (virtuelle Tasten bzw. virtuelle Schalter)
- 404 Linke obere Ecke der Bildfläche des VRD
- 405 Rechte obere Ecke der Bildfläche des VRD
- 406 Rechte untere Ecke der Bildfläche des VRD
- 407 Linke untere Ecke der Bildfläche des VRD
- 408 z. B. rechte Hand des Benutzers
- 501 Videokamera
- 502 Optischer bzw. akustischer Abstandssensor
- 503 Zentrale Steuerungseinheit (CCU)
- 504 Steuerungseinheit zur Steuerung des VRD
- 505 Steuerungseinheit für ein laserbetriebenes VRD
- 506 Perception und Verarbeitung der aufgenommenen Information durch den menschlichen Benutzer und Reaktion des Benutzers auf die veränderte Sachlage

#### Patentsprüche

1. Verfahren zur visuellen Darstellung von Objekten auf mindestens einem Sichtfeld (203), dadurch gekennzeichnet, dass mindestens ein dargestelltes Objekt mit Hilfe interaktiver Steuerbefehle eines Benutzers in seinen Eigenschaften und/oder Aktionen manipuliert, gesteuert bzw. beeinflusst werden und/oder die Möglichkeit einer Navigation des Benutzers in einer dargestellten Szene (204) gegeben ist, wobei die Erfassung und Verarbeitung der vom Benutzer eingegebenen Information mit Hilfe von Methoden der Signal- bzw. Mustererkennung erfolgt und somit keine zusätzlichen manuell bedienbaren, mechanischen bzw. berührungssensitiven Hardware-Vorrichtungen zur Eingabe von Steuerbefehlen benötigt werden.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass
  - a) die interaktiven Steuerbefehle des Benutzers mit Hilfe von Sensoren (502) und/oder Aufnahmegeräten (501) registriert werden,
  - b) die erfassten Eingabe-Daten einer Auswert- und Steuerungsvorrichtung (503) zugeführt und als Steuerbefehle interpretiert werden und
  - c) die Objekte entsprechend diesen Steuerbefehlen manipuliert bzw. beeinflusst werden oder/und eine Aktion ausgelöst wird
  - d) die manipulierten bzw. beeinflussten Objekte oder/und die ausgelöste Aktion auf einem Display (204) visualisiert bzw. akustisch und/oder optisch angezeigt werden.
3. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass es sich bei den zu

steuernden Objekten um real existierende Objekte in einer realen Umgebung handelt.

4. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass es sich bei den zu steuernden Objekten um nicht real existierende Objekte in einer virtuellen Umgebung eines computergesteuerten Modells handelt. 5

5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass es sich bei den interaktiven Steuerbefehlen eines Benutzers um akustische Signale, wie etwa Sprachkommandos, handelt. 10

6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass es sich bei der Vorrichtung (503) zur Interpretation von Steuerbefehlen um ein automatisches Erkennungssystem für Sprachsignale handelt. 15

7. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass es sich bei den interaktiven Steuerbefehlen eines Benutzers um optische Signale, wie etwa Körper-, Kopf-, Fuß-, Hand- und/oder Fingerbewegungen, also Gebärden, Gestik und/oder Mimik, handelt. 20

8. Verfahren nach einem der Ansprüche 1 bis 4 oder 7, dadurch gekennzeichnet, dass es sich bei den Sensoren (502) und/oder Aufnahmegeräten (501) zur Registrierung von optischen Signalen des Benutzers um Abstandssensoren und/oder Videokameras handelt. 25

9. Verfahren nach einem der Ansprüche 1 bis 4, 7 oder 8, dadurch gekennzeichnet, dass es sich bei der Vorrichtung zur Interpretation von Steuerbefehlen um ein automatisches Erkennungssystem für Bewegungssignale handelt. 30

10. Verfahren nach einem der Ansprüche 1 bis 9, dadurch gekennzeichnet, dass es sich bei der Ausgabevorrichtung 505 zur Darstellung einer virtuellen Umgebung um ein Gerät zur Erzeugung eines sogenannten "Virtual Retinal Displays" (VRD) handelt, bei dem ein virtuelles Display (203) auf die Netzhaut des Anwenders projiziert wird. 35

11. Verfahren nach einem der Ansprüche 1 bis 4 oder 7 bis 10, dadurch gekennzeichnet, dass Eingaben ausgelöst werden, wenn ein reales Objekt die Ebene des virtuellen Displays (401) an einer eingabesensitiven Fläche (403) schneidet bzw. berührt. 40

12. Verfahren nach einem der Ansprüche 1 bis 4 oder 7 bis 11, dadurch gekennzeichnet, dass die Ermittlung des Abstands zwischen der Ebene (303) des virtuellen Displays und der Netzhaut eines der Augen (304) des Benutzers bzw. die Begrenzung und räumliche Ausrichtung der Fläche (401) des virtuellen Displays durch die Erkennung von optischen Signalen des Benutzers mit Hilfe eines Abstandssensors (502) sowie einer nachgeschalteten Auswertelektronik (503) vorgenommen wird. 45 50

13. Verfahren nach einem der Ansprüche 1 bis 9, dadurch gekennzeichnet, dass es sich bei der Ausgabevorrichtung (505) zur Darstellung einer virtuellen Umgebung um ein sogenanntes "Head-Mounted Display" (HMD) handelt. 55

---

Hierzu 4 Seite(n) Zeichnungen

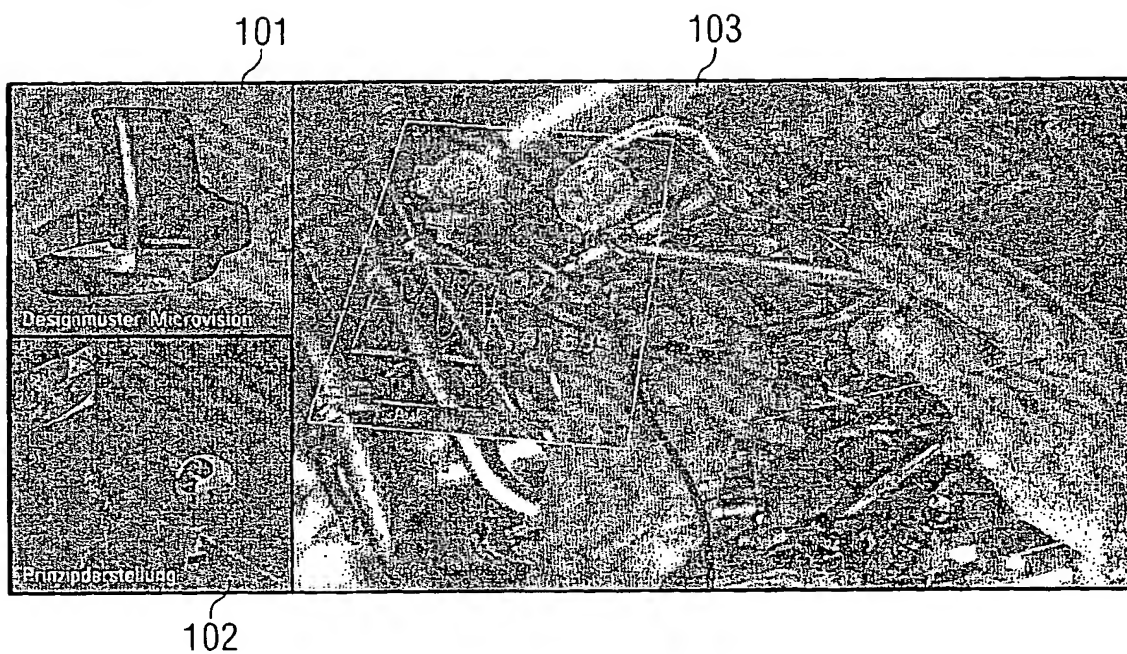
---

60

65

- Leerseite -

FIG 1



BEST AVAILABLE COPY

FIG 2

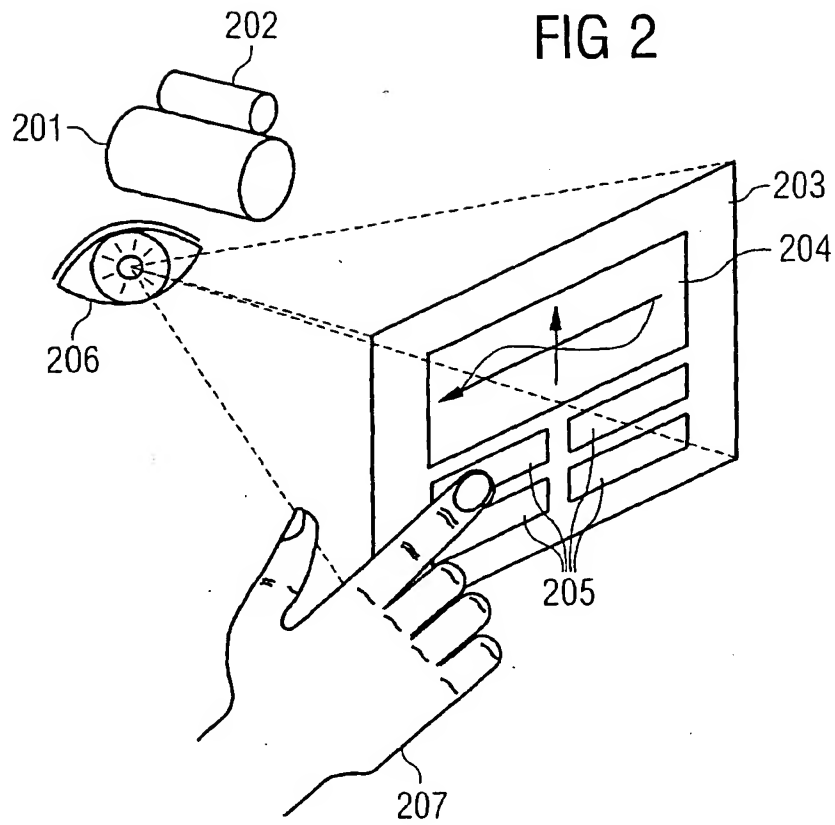


FIG 3

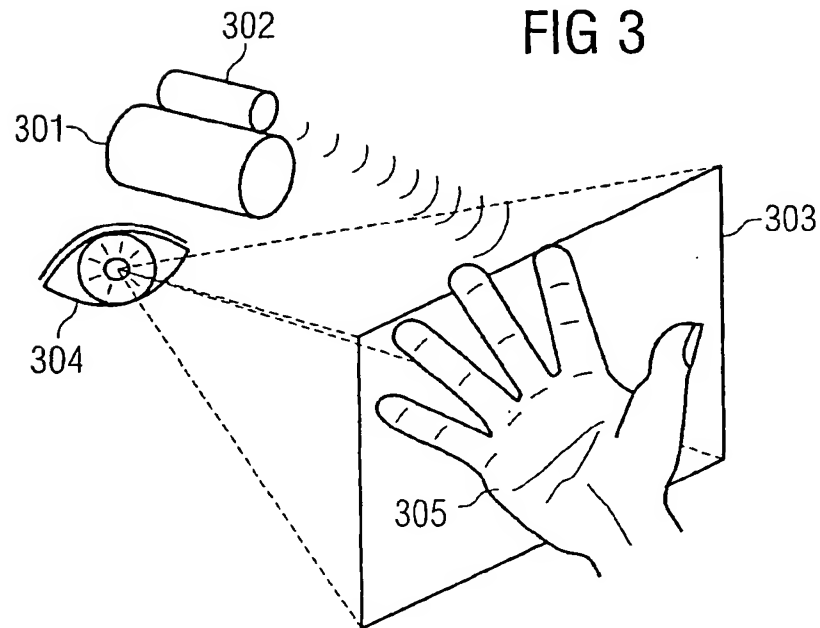




FIG 4

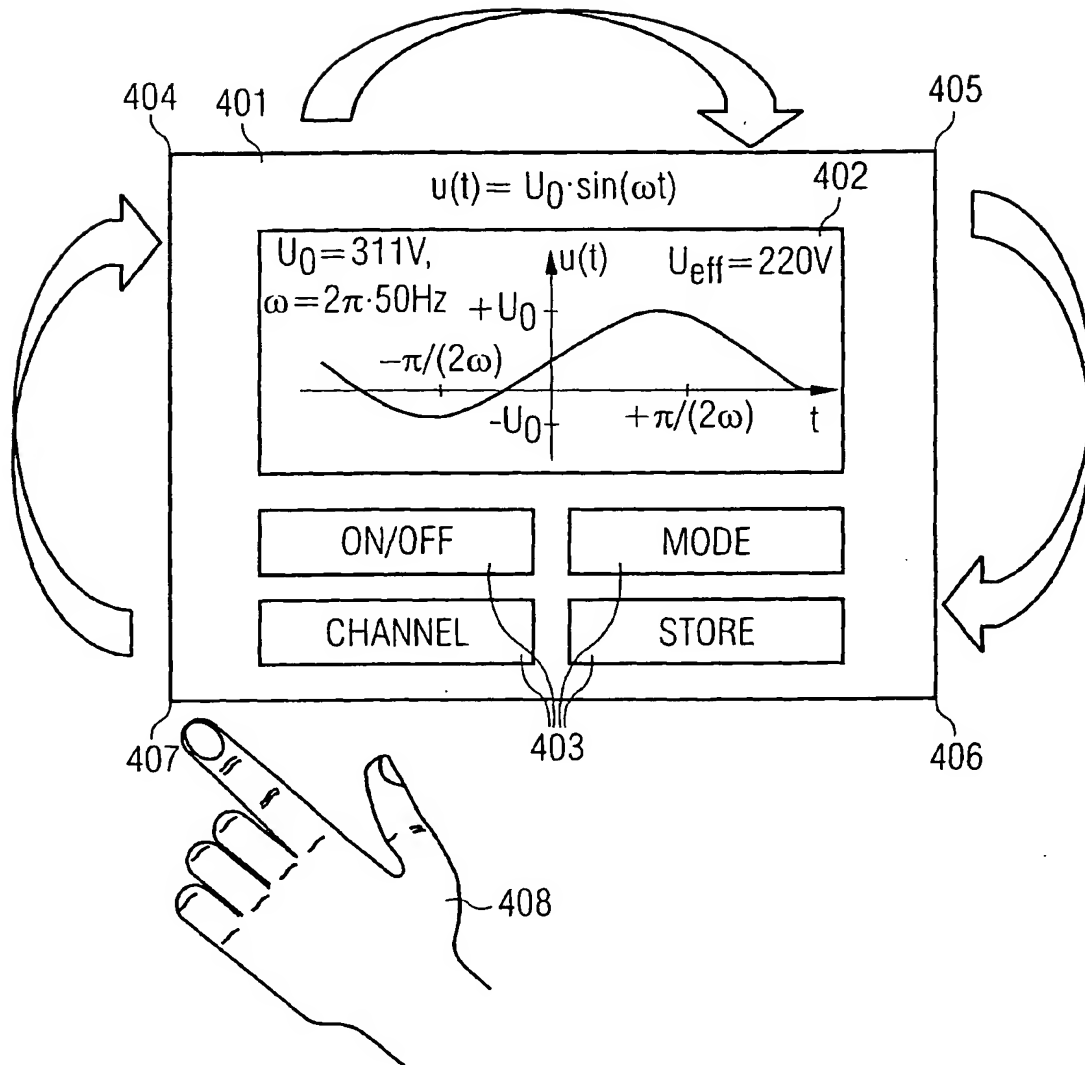


FIG 5

